Nesting Cohort Affiliation of Stranded *Caretta caretta* Recovered in Georgia

Final Report to Georgia Department of Natural Resources

Prepared by

Anna L. Bass, Shiao-Mei Chow, and Brian W. Bowen
Department of Fisheries and Aquatic Sciences
7922 NW 71st Street
Gainesville, Florida 32653

Introduction

Recent research has demonstrated that most sea turtle nesting colonies are genetically distinct in terms of mitochondrial (mt) DNA haplotype frequency shifts. This finding allows the possibility of using mtDNA data to identify rookery cohorts on feeding grounds (Bass et al., 1998). Utilizing existing databases (Encalada et al., 1998 and unpublished data) and molecular techniques, tissue samples from stranded marine turtles can be useful to estimate the origin of animals inhabiting coastal waters. The focus of this project is on loggerhead turtles, *Caretta caretta*, in Georgia with the goal of identifying the origin of turtles stranded along the coast.

Methods

Tissue samples were collected by Sea Turtle Stranding Network volunteers and placed in 15 ml of saturated salt preservation buffer. Samples were then transferred to the University of Florida for analysis. Standard phenol/chloroform DNA isolation protocols were conducted on the tissue samples and a 380 bp fragment of the mitochondrial DNA control region was amplified using primers designed for sea turtles (Allard et al., 1994; Norman et al., 1994). Individual fragments were sequenced and compared to known *Caretta caretta* nesting beach haplotypes. Individuals were then assigned a haplotype based on designations from Encalada et al. (1998).

To test for statistical differences among haplotype frequencies at rookeries and the stranding cohort, chi-square analyses were performed with the program CHIRXC (Zaykin and Pudovkin, 1993) and probabilities were generated using a Monte Carlo randomization procedure (Roff and Bentzen, 1989). To correct for simultaneous tests, the sequential Bonferroni technique (Rice, 1989) was applied to the probabilities generated by CHIRXC.

Maximum likelihood (ML) analysis for mixed stock identification (Grant et al., 1980) was used to estimate the contributions of nesting populations to coastal regions in Georgia. This method estimates the most likely contributions of source populations based on the haplotype frequencies in the source populations. The maximum likelihood programs, UCON and GIRLSEM were used (Masuda et al., 1991). UCON differs from GIRLSEM in assuming that all source populations have been identified. As a starting point in ML iterations, it was assumed that all source populations had an equal probability of contributing (i.e. population size, distance from the foraging location, etc. did not have an impact in the percentage of animals recruiting to a particular area). The standard deviation was estimated using an infinitesimal jackknife procedure in the ML programs.

Results and Discussion

Of the 111 samples from stranded turtles provided, only 10 did not produce readable sequence (Appendix 1). The most common haplotype in the sample of 101 that did provide readable sequence was haplotype A which is found at nesting locations along the southeastern United States (Table 1). The second most common haplotype was haplotype B which is found in the south Florida, Mexico and Greek nesting populations. Haplotype B has also been observed in northwest Florida (NWFL) and in NEFL-NC, but in low frequency (Encalada et al., 1998). Haplotype C has been observed in low frequency in multiple nesting locations (see Table 1). Haplotype G has been observed in low frequency in NWFL and southwest Florida (a part of the south Florida (SFL) management unit). The remaining identified haplotypes, K, M, and N, have been observed only in foraging populations in the North Atlantic (Bolten et al., 1998) or stranded along the North Atlantic coast (Rankin-Baransky et al., 1998).

Table 1. Haplotypes identified in the sample of Georgia loggerhead turtle strandings and the locations where these haplotypes have been observed.

Haplotype	Number of Individuals	Location
		Nesting[1]
A	51	NWFL, SFL, NEFL-NC
B	35	NWFL, SFL, NEFL-NC, Mexico, Greece
C	6	NWFL, SFL, Mexico
G	1	NWFL, SFL
	(n = 93)	*Foraging*[2] *or Stranded*[3]
K	1	Madeira[2]
M	1	Azores[2]
N	6	Azores[2], Madeira[2], North Atlantic Coast[3]
	Total = 101	

[1] Abbreviations according to Encalada et al. (1998): NWFL = Northwest Florida; SFL = Southwest and southeast Florida; NEFL-NC = Northeast Florida to North Carolina.
[2] From Bolten et al. (1998)
[3] From Rankin-Baransky et al. (1998)

The results of the X^2 analyses indicated that the stranded cohorts from Georgia were significantly different from all nesting locations at the 0.05 level except for the comparison to the SFL management unit ($X^2 = 3.619$). There was a 50% probability that the haplotype frequencies of the stranded cohorts are the same

as the haplotype frequencies of the SFL management unit indicating that the SFL management unit may be a major contributor.

The results of maximum likelihood (ML) analysis are shown in Table 2. Because eight (of 101) individuals had haplotypes not observed in a nesting population, only 93 individuals were included. These analyses agree with the indications of the X^2 test that the South Florida source population is the primary contributor to the stranded cohort in Georgia. The estimates for the NWFL, and NEFL-NC source populations are more difficult to interpret due to the large standard deviations.

Table 2. Maximum likelihood estimates of contribution of source populations to stranded cohorts of loggerheads recovered in Georgia. Estimates and standard deviations were generated using the program UCON. Two separate analyses were conducted: 1) including Greece as a possible source population and 2) not including Greece as a possible source population.

Source Population	+ Greece	- Greece
NWFL	0.2029 (± 0.2324)	0.1588 (± 0.1762)
SFL	0.7331 (± 0.1148)	0.7391 (± 0.1106)
NEFL-NC	0.0624 (± 0.1891)	0.1019 (± 0.1421)
Mexico	0	0
Brazil	0	0
Greece	0.0015 (± 0.0003)	N/A

Because of the large standard deviations associated with the estimates shown in Table 2, an investigation was conducted regarding the possible source of the error(s) resulting from sampling of the stocks (source populations) and/or mixture (stranded cohorts). For the following analyses it was assumed that all source populations had been identified so that GIRLSEM could be used and bootstrapping of stocks and mixture could be conducted. The actual number of individuals carrying haplotype C in the mixture was greater than that found within the identified stocks. For example, NWFL and SFL both have two individuals with haplotype C. The occurrence of haplotype C in 6 individuals in the mixture may result in an overestimate of the actual contribution by NWFL to the stranded cohorts [this has been observed in simulation studies involving salmon (Pella and Milner, 1987)]. Statistical resampling (bootstrapping) of both the nesting samples and the strandings indicated that the source of error was due to the high frequency of 'C' individuals in the strandings. When three 'C' individuals are removed, decreasing the total number of 'C' individuals in the strandings to 3, the standard deviations decrease and the estimate of the contribution by the NWFL population makes more sense in terms of the size of the nesting population and proximity to Georgian waters (Table 3).

Table 3. Maximum likelihood estimates of contribution of source populations to stranded cohorts of loggerheads recovered in Georgia. Estimates and standard deviations were generated using the program GIRLSEM. To estimate the impact of a large number of individuals possessing haplotype C, three of the six individuals with this haplotype were removed from the sample of stranded cohorts (n=90).

Source Population	+ Greece	- Greece
NWFL	0.0172 (±0.0683)	0.0171 (±0.0680)
SFL	0.7643 (±0.0973)	0.7644 (±0.0973)
NEFL-NC	0.2183 (±0.1059)	0.2184 (±0.1058)
Mexico	0	0
Brasil	0	0
Greece	0	N/A

If we assume that NWFL is not contributing individuals at detectable levels to the stranded cohort, then this source population can be completely removed from the analysis. We have demonstrated that the effect of large numbers of underrepresented haplotypes in the mixture may result in an overestimate of a stock's contribution; therefore, the sample size was limited to n=90 (Table 4). As with the majority of these analyses the major contributors appear to be the SFL and NEFL-NC populations.

Table 4. Maximum likelihood estimates of contribution of source populations to stranded cohorts of loggerheads recovered in Georgia. Estimates and standard deviations were generated using the program GIRLSEM. Sample size was once again held to n=90. Due to the low contribution of NWFL as determined by previous analyses, it was removed from this analysis.

Source Population	+ Greece	- Greece
SFL	0.7690 (±0.0948)	0.7693 (±0.0947)
NEFL-NC	0.2306 (±0.0948)	0.2306 (±0.0947)
Mexico	0	0
Brasil	0	0
Greece	0.0003 (±0.0003)	N/A

Conclusions

A primary question motivating this analysis is whether stranded marine turtles can be used to estimate the origin of live turtles inhabiting coastal regions of Georgia. If stranded turtles are a random sample of live turtles in the area, then

maximum likelihood analysis of mtDNA sequences may provide robust estimates of foraging ground composition.

It appears that the major contributors to the coastal waters of Georgia are the SFL and NEFL-NC nesting populations or management units. This is not surprising due to the proximity and size of these nesting populations. Several recommendations arise from the results of these analyses and may also be considered caveats in discussing the composition of the stranded cohorts in Georgia. The haplotypes C and G are found at low frequency in several of the potential source populations (NWFL, SFL and Mexico; see Encalada et al. 1998). The appearance of 6 'C' individuals and one 'G' individual is a qualitative indicator that the NWFL and Mexico populations are contributing at low levels to the stranded cohorts in Georgia. An increase in the sample size of stranded individuals may provide evidence of other haplotypes that are endemic to the Mexican population. Investigations into the stock structure of fish has shown that increases in sample size can provide several benefits, such as a reduction in the standard errors of frequency estimates for common haplotypes and an increase in the probability of detecting 'endemic' haplotypes which in turn increases resolution (Epifanio et al., 1995).

In conclusion we regard these estimates as provisional in determining all potential contributors and their respective proportions. It is apparent that both the SFL and NEFL-NC nesting populations contribute approximately 75% and 20% of the individuals respectively, but at this point we cannot exclude other source populations with finality. We feel it is likely that other sources (NWFL, Mexico) contribute to the Georgia feeding population at low frequency. Further investigation may very well provide evidence of contributions from other nesting populations.

Literature Cited

Allard, M.W., M.M. Miyamoto, K.A. Bjorndal, A.B. Bolten, and B.W. Bowen. 1994. Support for natal homing in green turtles from mitochondrial DNA sequences. Copeia 1994:34-41

Bass, A. L., C. J. Lagueux, and B. W. Bowen. 1998. Origin of green turtles, *Chelonia mydas*, at "sleeping rocks" off the northeast coast of Nicaragua. Copeia *IN PRESS*.

Bolten, A.B., K.A. Bjorndal, H.R. Martins, T. Dellinger, M.J. Biscoito, S.E. Encalada, and B.W. Bowen. 1998. Transatlantic developmental migrations of loggerhead sea turtles demonstrated by mtDNA sequence analysis. Ecological Applications 8(1):1-7.

Encalada, S.E., K.A. Bjorndal, A.B. Bolten, J.C. Zurita, B. Schroeder, E. Possardt, C.J. Sears, B.W. Bowen. 1998. Population structure of loggerhead turtle (*Caretta caretta*) nesting colonies in the Atlantic and Mediterranean as inferred from mitochondrial DNA control region sequences. Marine Biology 130:567-575.

Epifanio, J.M., P.E. Smouse, C.J. Kobak, and B.L. Brown. 1995. Mitochondrial DNA divergence among populations of American shad (*Alosa sapidissima*): how much variation is enough for mixed-stock analysis? Canadian Journal of Fisheries and Aquatic Sciences 52:1688-1702.

Grant, W.S., G.B. Milner, P. Krasnowski, and F.M. Utter. 1980. Use of biochemical genetic variants for identification of sockeye salmon (*Oncorrhynchus nerka*) stocks in Cook Inlet, Alaska. Canadian Journal of Fisheries and Aquatic Sciences 37:1236-1247.

Masuda, M., S. Nelson, and J. Pella. 1991. User manual for GIRLSEM, GIRLSYM, AND CONSQRT. USA-DOC-NOAA-NMFS. Programs and user manual available from US-Canada Salmon Program, 11305 Glacier Hwy., Juneau, Alaska 99801 USA.

Norman, J.A., C. Moritz, C.J. Limpus. 1994. Mitochondrial DNA control region polymorphisms: Genetic markers for ecological studies of marine turtles. Molecular Ecology 3:363-373.

Pella, J.J. and G.B. Milner. 1987. Use of genetic marks in stock composition analysis. Pp. 247-276 *in* N. Ryman and F. Utter, eds. Population genetics and fisheries management. Univ. of Washington Press, Seattle.

Rankin-Baransky, K., C.J. Williams, A.L. Bass, B.W. Bowen, and J.R. Spotila. 1998. Origin of loggerhead turtle strandings in the northwest Atlantic as determined by mtDNA analysis. *IN PREP.*

Rice, W.R. 1989. Analyzing tables of statistical tests. Evolution 43:223-225.

Roff, D.A. and P. Bentzen. 1989. The statistical analysis of mitochondrial DNA polymorphisms: Chi-square and the problem of small samples. Molecular Biology and Evolution 6:539-545.

Zaykin, D.V., and A.I. Pudovkin. 1993. Two programs to estimate significance of X2 values using pseudo-probability tests. Journal of Heredity 84:152.